



DØNote 5275-CONF

## A Search for $ZH(\rightarrow l^+l^-b\bar{b})$ Production with the DØ Detector in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration  
URL <http://www-d0.fnal.gov>  
(Dated: November 13, 2006)

We report on a search for a Higgs boson produced association with a  $Z$  boson at DØ in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV at the Fermilab Tevatron collider. Events containing  $Z \rightarrow e^+e^-$  or  $\mu^+\mu^-$  and two  $b$ -tagged jets are considered, and the integrated luminosity is  $920 \text{ pb}^{-1}$  ( $840 \text{ pb}^{-1}$ ) in the dielectron (dimuon) channel. Good agreement between data and the expected background has been observed. The upper limits on the  $ZH$  production cross section are set for Higgs masses between 105 and 155 GeV.

*Preliminary Results for Summer 2006 Conferences*

## I. INTRODUCTION

One of the most sensitive search channels at the Tevatron for a Standard Model Higgs boson with a mass below approximately 140 GeV is the associated production of a Higgs boson with a  $Z$  boson, and the Higgs decays to  $b\bar{b}$ . In this note we present a search for the  $ZH$  production in  $l^+l^-b\bar{b}$  final states. The product of cross section and branching fraction ( $\sigma(p\bar{p} \rightarrow ZH) \times Br(H \rightarrow b\bar{b})$ ) is predicted to be 0.12–0.006 pb for a Standard Model Higgs boson with a mass between 105 and 155 GeV.

The  $Z$  boson is reconstructed and identified from a pair of high  $p_T$  leptons with an invariant mass constraint. Events are required to have at least two  $b$ -tagged jets. We then search for a  $H \rightarrow b\bar{b}$  resonance in the dijet mass distribution. The dominant backgrounds result from the associated production of a  $Z$  boson with jets, among which the  $Zb\bar{b}$  production is an irreducible background. The other main backgrounds are  $t\bar{t}$ ,  $WZ$ ,  $ZZ$ , and multijet production from QCD processes. In order to isolate these background sources, an efficient  $b$ -tagging algorithm and a good dijet mass resolution are essential.

## II. DATA SAMPLE AND EVENT SELECTION

The DØ detector has a central-tracking system, consisting of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet [1], with tracking and vertexing at pseudorapidities  $|\eta| < 3$  and  $|\eta| < 2.5$ , respectively. A liquid-argon and uranium calorimeter has a central section (CC) covering  $|\eta|$  up to  $\approx 1.1$ , and two end calorimeters (EC) that extend coverage to  $|\eta| \approx 4.2$  [2]. An outer muon system, at  $|\eta| < 2$ , consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroids, followed by two similar layers after the toroids.

The analyses are based on data which have been taken between April 2002 and May 2006. No trigger requirements were made, in order to retain the highest possible efficiency for signal. The integrated luminosities were found to be  $920 \text{ pb}^{-1}$  and  $840 \text{ pb}^{-1}$  for the dielectron and the dimuon channels respectively, after the requirement of good data quality.

In the subsections below, we describe event selection specific to each channel. The selection of hadronic jets is common to both analyses. A jet is reconstructed using the RunII cone algorithm with  $\Delta R = 0.5$  [3]. The jet must have  $p_T > 15$  GeV after the jet energy scale correction and a pseudorapidity of  $|\eta| < 2.5$ . In the simulation, jets are smeared to account for the difference in reconstruction efficiency with the data. All corrections to the jets and leptons are propagated in the computation of the transverse missing energy  $\cancel{E}_T$ .

### A. Dielectron Sample Selection

Events are required to have at least two electron candidates, with  $p_T > 15$  GeV,  $|\eta_{\text{detector}}| < 1.1$  or  $1.5 < |\eta_{\text{detector}}| < 2.5$ , and satisfying electron shower shape criteria and a match with a central track. In the simulation, the EM identification efficiency is corrected to the data one as a function of  $\eta$  for each electron.

A good  $Z$  candidate is required, reconstructed from two electrons. The reconstructed  $Z$  mass must be  $> 65$  and  $< 115$  GeV. At least two jets are required in each event with each jet being away from both good electrons by  $\Delta R > 0.5$ .

### B. Dimuon Sample Selection

Events are required to have at least two jets with  $p_T > 15$  GeV and two Loose quality muons, each matched with a central track [4], and  $p_T > 10$  GeV. The muon  $p_T$  in data is corrected using the beam spot position for each run, if the muon track has no SMT hits.

A good  $Z$  candidate is required, reconstructed from muons with  $p_T > 15$  GeV. The reconstructed  $Z$  mass must be  $> 70$  and  $< 110$  GeV. The two muons which form the  $Z$  are together required to have *product scaled isolation*  $< 0.01$ . The product scaled isolation is the product of the scaled isolation variables for each of the two muons from the  $Z$ . Scaled isolation is defined as  $(\text{track } p_T + \text{calorimeter } p_T) / \text{muon } p_T$ . The track  $p_T$  is the sum of all other tracks in a cone of  $\Delta R < 0.5$  around the muon track (and not including the muon track). The calorimeter  $p_T$  is the sum of all calorimeter energy in a hollow cone around the muon from  $0.1 < \Delta R < 0.4$  [4].

The reconstructed  $Z$   $p_T$  must be greater than 20 GeV. This improves the expected signal significance since the Higgs signal peaks at about 45 GeV, whereas the background peaks around 20 GeV.

### C. $b$ -jet Tagging

For both analyses, at least two jets must be  $b$ -tagged using a Neural Net based  $b$ -tagging algorithm [5]. In simulation, the tag-rate-function (TRF) for data was used to predict the probability that a jet of a given flavor would be tagged. The NN  $b$ -tagging requirement was adjusted to accept a 4% light-jet fake rate, providing a 72%  $b$ -tagging efficiency in the central region ( $|\eta| < 1.5$ ), for  $p_T > 15$  GeV.

### III. QCD BACKGROUND

In the dielectron channel analysis, the size of the QCD contribution is estimated by fitting the dielectron invariant mass by a Breit-Wigner function convoluted with a Gaussian for the  $Z$  peak ( $N_Z$ ), and an exponential shape for QCD and Drell-Yan processes ( $N_{QCD} + N_{DY}$ ). The fraction of Drell-Yan events is determined from the simulation. Out of 2900  $Z + \geq 2$  jets events within  $Z$  mass range, 102 events are estimated to be from QCD background. The shape of the QCD background is obtained from data where the QCD contribution is enhanced by inverting the electron Likelihood requirement for electron identification (which also removes the track requirement). This sample does not have a  $Z$  peak in the dielectron mass distribution, and is used to predict the QCD background in kinematical distributions.

In the dimuon channel analysis, the QCD fraction is determined by fitting an exponential plus Gaussian to the data. The Drell-Yan (DY) fraction is determined from a fit to simulation to be 4.8% and will be subtracted from the fits to data. The fits are performed to the data in events with at least two jets (since the QCD fraction was observed to depend on the jet multiplicity). A separate fit is performed for events with 0, 1, and at least 2  $b$ -tagged jets. The QCD fractions in the 0, 1, and 2  $b$ -tagged categories are found to be 1.0%, 0.5%, and 21%, respectively. The significant increase in the relative amount of QCD in double  $b$ -tagged data is assumed to be due to the presence of  $b\bar{b}$  events, which can frequently produce the searched signature since the muon(s) from semileptonic  $b$ -decay may pass the selection requirements.

### IV. SIMULATED EVENT SAMPLES

Using the CTEQ6L1 [6] leading-order parton distribution function, the following physics processes are simulated to estimate the signal acceptance and the number of background events:  $Z(\rightarrow l^+l^-)H(\rightarrow b\bar{b})$  by PYTHIA,  $Z(\rightarrow l^+l^-)jj$  including  $Z(\rightarrow l^+l^-)cc$  by ALPGEN,  $Z(\rightarrow l^+l^-)b\bar{b}$  by ALPGEN,  $t\bar{t} \rightarrow l^+\nu b l^-\bar{\nu}\bar{b}$  and  $t\bar{t} \rightarrow bbj\bar{j}l\nu$  by ALPGEN, inclusive  $ZZ$  and  $WZ$  by PYTHIA. The samples generated by ALPGEN are interfaced with PYTHIA for parton showering and hadronization. All the samples were run through the full GEANT detector simulation, digitization, and reconstruction.

The signal cross-sections, as well as those for  $t\bar{t}$ ,  $WZ$ , and  $ZZ$  are taken from MCFM [7], which is NLO. For the ALPGEN samples, a matching procedure (MLM) was used so as to not double count the radiation of additional jets between ALPGEN and PYTHIA. In addition, NLO  $k$ -factors of 1.23 and 1.35 were applied to the  $Z$ +light and  $Z+b\bar{b}$  samples, respectively, to account for the NLO cross-sections of these process (as calculated with MCFM) as compared to the LO cross-sections from ALPGEN. The scale factor for  $Z+b\bar{b}$  is considered to be a rough approximation, good to  $\pm 50\%$ , while a much more thorough correction or solution is currently being evaluated.

### V. ANALYSIS

In the dielectron analysis, simulation samples are scaled up by 10% to agree with number of observed data under  $Z$  mass range (65–115 GeV). Equivalently the effective integrated luminosity is  $920 \times 1.1 = 1010 \text{ pb}^{-1}$ , which includes possible trigger inefficiency and electron inefficiencies. After requiring two electrons and two jets in the events, and  $M_{ee}$  between 65 GeV and 115 GeV, 2900 events are observed in data, while 2891 events are expected from QCD + simulation background. Since the  $Z p_T$  distribution is poorly modeled in  $Z$  + light jet simulation, the  $Z$  + light jet simulated samples are reweighted according to  $Z p_T$  before  $b$ -tagging. After such reweighting, Figure 1 shows the  $Z p_T$  and  $M_{ee}$  distributions for data and background.

In the dimuon analysis, all  $Z$  + light jets simulation is scaled to match the  $Z$  peak in data, after requiring only a good reconstructed  $Z$  (see criteria below), with no criteria on the jets in the events. Given the measured integrated luminosity of  $840 \text{ pb}^{-1}$ , the  $Z$  peak in data is found to have a cross-section which is 87% of the 255 pb expected at NNLO. This scale factor (0.87) is then applied to all other simulation samples as well. The effective integrated luminosity is  $840 \times 0.87 = 730 \text{ pb}^{-1}$ , which contains trigger inefficiency and muon ID inefficiencies.

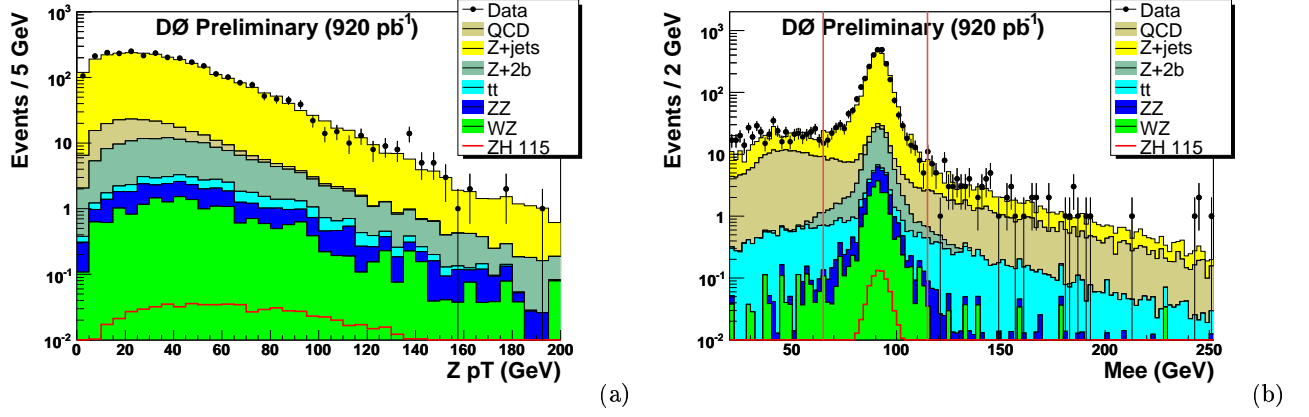


FIG. 1: After  $Z p_T$  reweighting for  $Z + \text{light jet}$  simulation and before  $b$ -tagging, the distributions for (a)  $Z p_T$ , and (b) dielectron invariant mass.

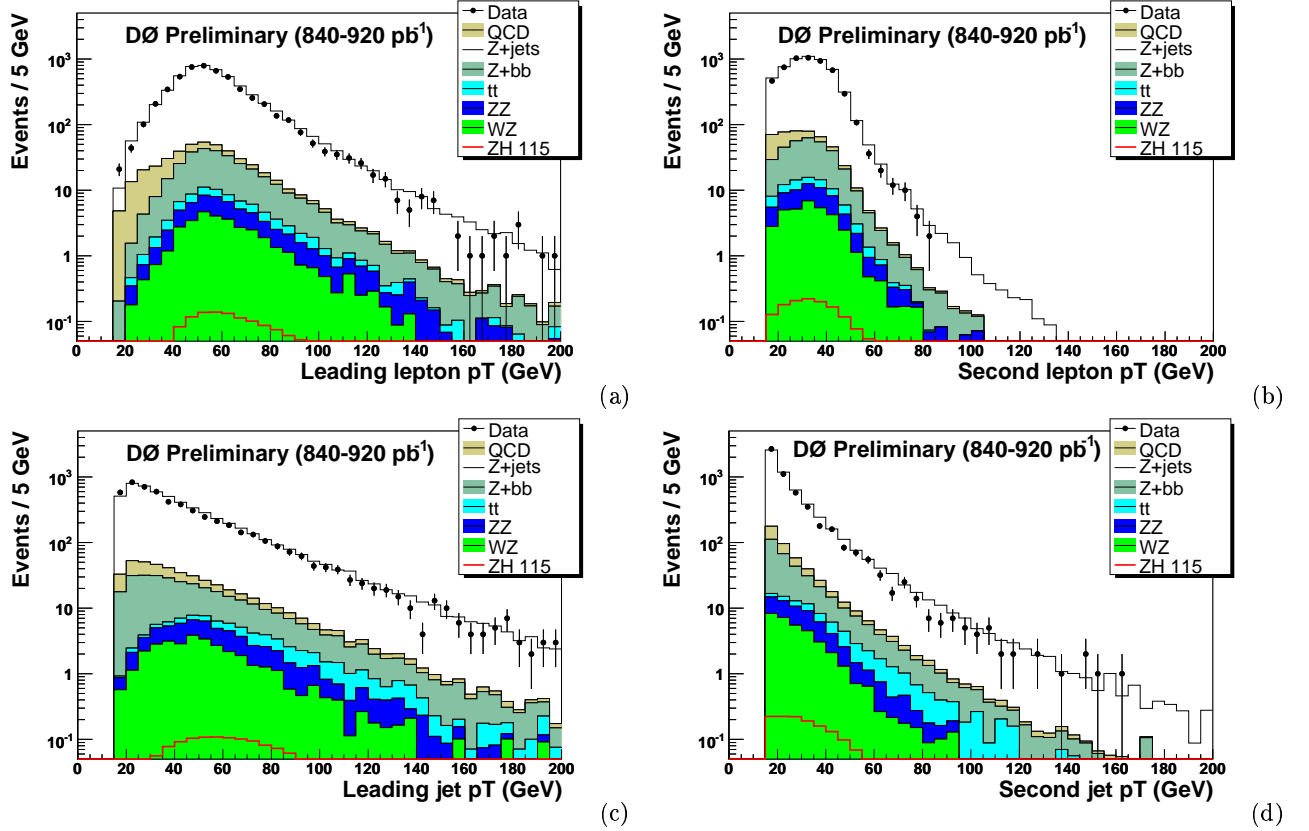


FIG. 2: The lepton and jet  $p_T$  distributions for  $Z(\rightarrow l^+l^-) + \geq 2$  jets events: (a) leading lepton  $p_T$ , (b) second lepton  $p_T$ , (c) leading jet  $p_T$ , (d) second jet  $p_T$ .

Figure 2 shows the  $p_T$  distributions for the leading two leptons and two jets in the  $Z + \geq 2$  jets events, where the dielectron and dimuon samples are combined.

### A. Search Windows

Since the Higgs is decaying to  $b\bar{b}$ , we attempt to isolate the signal by selecting events where the leading jets form an invariant mass close to that expected for a given Higgs mass. The actual value of the invariant mass reconstructed in full simulation is less than the hypothetical Higgs mass due to final-state radiation of particles and jets outside

TABLE I: The dijet mass window to search for Higgs signal.

Higgs mass (GeV)	dielectron sample (GeV)	dimuon sample (GeV)
105	56.2-115.2	66.0-124.7
115	63.7-125.6	72.1-136.6
125	67.2-134.1	77.6-148.4
135	71.9-146.8	84.3-159.9
145	76.7-157.3	89.8-171.7
155	82.2-167.5	96.3-182.6

TABLE II: The systematic uncertainties for background and signal simulation samples.

Sample	JES (%)	$b$ -tagging & taggability (%) $\geq 2$ $b$ -tag	cross section (%)	lepton-ID (%)
$Z + (udscg)$	7.0	8.4	15	4
$Z + 2b$	5.8	7.3	50	4
$t\bar{t} \rightarrow l^+l^-$	0.8	7.5	8	4
$t\bar{t} \rightarrow l\nu$	0.8	6.8	8	4
$WZ$	2.9	8.5	7	4
$ZZ$	1.6	7.2	6	4
$ZH$ 105	1.8	7.3		4
$ZH$ 115	1.2	7.4		4
$ZH$ 125	1.1	7.4		4
$ZH$ 135	0.8	7.5		4
$ZH$ 145	0.8	7.6		4
$ZH$ 155	0.6	7.7		4

the reconstructed jet cones. The leading- $p_T$  dijet mass distribution in each Higgs simulation samples is fit to a Gaussian, in the double  $b$ -tagged channel. The search window extends from mean $-1.5 \times$ width to mean $+1.5 \times$ width in the dielectron analysis, and from mean $-1.0 \times$ width to mean $+2.0 \times$ width in the dimuon channel. The search windows are summarized in Table I.

## B. Systematic Uncertainties

In the dielectron analysis, the QCD background has a 30% uncertainty on normalization factor, obtained from fitting various jet multiplicity samples.

For background and signal simulation samples, the uncertainties from various sources are estimated and listed in Table II. The uncertainties from Jet Energy Scale,  $b$ -tagging and taggability are estimated by varying JES and  $b$ -tagging TRF by  $\pm 1$  standard deviation. The cross sections uncertainties are taken from the previous  $ZH \rightarrow e^+e^-bb$  analysis [8], except for  $Z+2b$  sample. The uncertainties of EMID efficiency or scale factor for dielectron events is 4%.

The total uncertainty is calculated by adding the uncertainties linearly for the same error source and adding them in quadrature for different error sources. The total background uncertainty is 35%, and the signal efficiency uncertainty is 9% for double  $b$ -tagged samples.

To check the 50% uncertainty assigned to  $Z+2b$  sample cross section can cover the shape difference between data and ALPGEN simulation, data and simulated samples are reprocessed using a tighter  $b$ -tagging operation point and the single  $b$ -tagged events are selected. In this way, light-jet backgrounds were suppressed while we still have not too few  $Z+2b$  events to look at. The total backgrounds agree well with data in the dijet invariant mass distribution. We have 96 events observed in data, and 96.3 events expected from backgrounds, among which 41.3 events are from  $Z+2b$ . By counting outside of the  $ZH$  search window for 115 GeV Higgs (about 60-120 GeV in dijet mass range), 51 events are observed in data, and  $55.3 \pm 7.7$  events are expected from backgrounds, among which 23.1 events are expected from  $Z+2b$ . Thus the statistical fluctuation and the systematic uncertainty of the total background are within the 50% uncertainty of  $Z+2b$  simulation.

TABLE III: The data, backgrounds, and expected SM signal after 0, 1, and  $\geq 2$   $b$ -tagged jets are required, for a representative dijet mass range (60–110 GeV) in dielectron channel and (70–110 GeV) in dimuon channel.

Sample	dielectron channel			dimuon channel		
	0 tag	1 tag	$\geq 2$ tag	0 tag	1 tag	$\geq 2$ tag
Data	832	126	8	647	99	10
$ZH$ 115	$0.111 \pm 0.004$	$0.243 \pm 0.011$	$0.169 \pm 0.014$	$0.11 \pm 0.004$	$0.21 \pm 0.009$	$0.14 \pm 0.012$
Total Bgnd.	$808 \pm 134$	$111 \pm 22$	$9.8 \pm 3.4$	$671 \pm 111$	$88 \pm 18$	$8.7 \pm 3.1$
QCD	28.5	4.42	0.25	6.71	0.44	1.85
$Z + (udscg)$	740	82.4	3.69	661.97	75.45	3.70
$Z + 2b$	19.3	17.8	3.78	21.23	16.02	3.35
$t\bar{t}$	1.02	1.98	1.25	0.92	1.75	1.22
$WZ$	10.93	2.20	0.105	6.87	1.35	0.07
$ZZ$	8.67	2.32	0.750	6.60	2.02	0.65

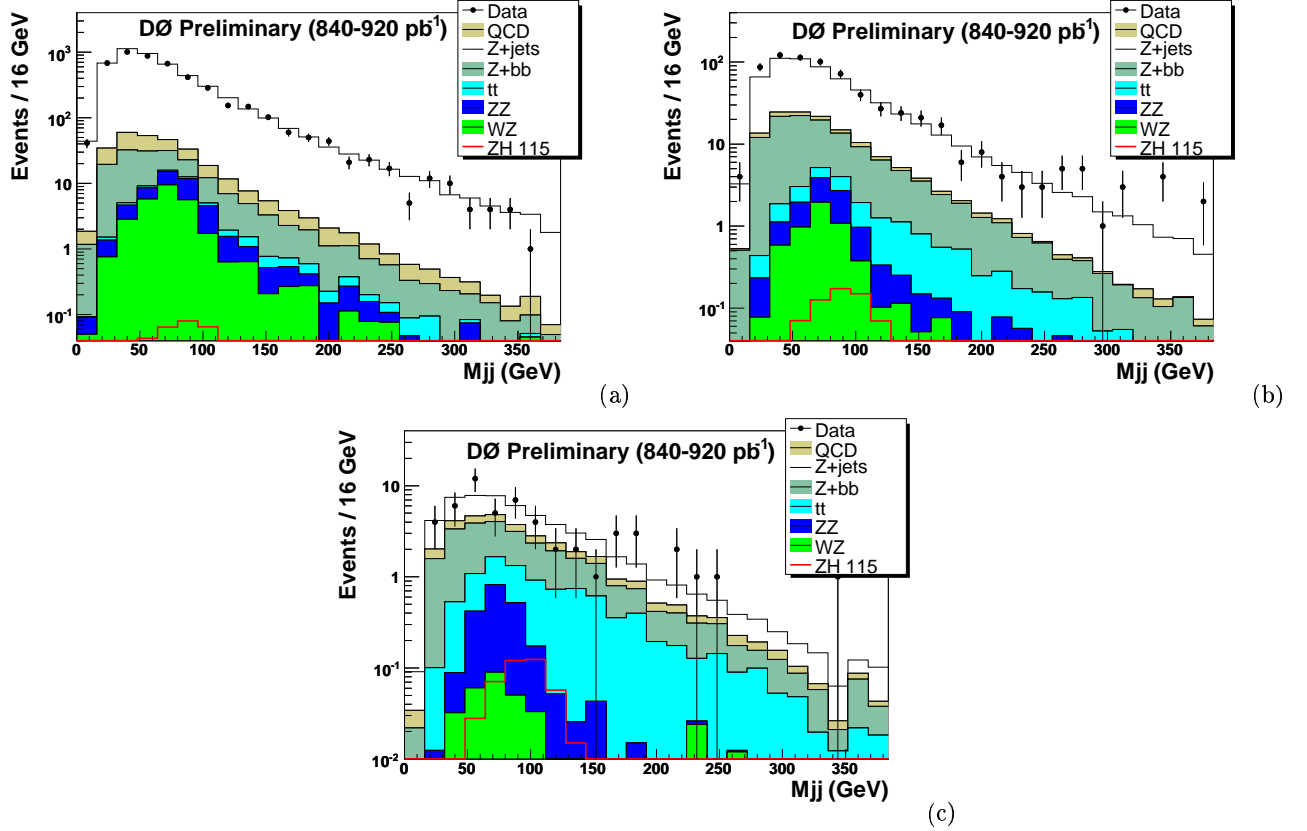


FIG. 3: The leading- $p_T$  dijet invariant mass distribution in the events: (a) 0  $b$ -tagged jets, (b) 1  $b$ -tagged jets, (c)  $\geq 2$   $b$ -tagged jets. The dielectron and dimuon samples are combined.

### C. Results

The total amount of data, various backgrounds, and expected signal are shown in Table III after 0, 1, and at least 2  $b$ -tagged jets are required, for a representative dijet mass range (60–110 GeV) in dielectron channel and (70–110 GeV) in dimuon channel. Figure 3 shows the invariant mass distribution of the two leading- $p_T$  jets, after requiring 0, 1, and at least 2  $b$ -tagged jets in the events, for the combined dielectron and dimuon samples.

TABLE IV: For  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$  inclusive double  $b$ -tagged samples, the number of events of data, total background, and expected signal in each Higgs mass window.

$m_H$ (GeV)	dielectron channel			dimuon channel		
	Data	Bgnd.	Sig.	Data	Bgnd.	Sig.
105	12	$11.6 \pm 4.1$	0.253	11	$10.1 \pm 3.5$	0.21
115	7	$10.7 \pm 3.7$	0.198	11	$9.8 \pm 3.4$	0.17
125	7	$10.6 \pm 3.7$	0.132	11	$10.4 \pm 3.6$	0.12
135	6	$10.5 \pm 3.7$	0.088	9	$9.0 \pm 3.1$	0.07
145	7	$10.3 \pm 3.6$	0.043	9	$8.1 \pm 2.8$	0.04
155	8	$9.8 \pm 3.4$	0.018	7	$7.4 \pm 2.6$	0.02

TABLE V: The expected and observed  $ZH$  cross section limits for each Higgs mass, derived from the combination of dielectron and dimuon samples.

$m_H$ (GeV)	105	115	125	135	145	155
Observed limit (pb)	3.31	2.72	2.65	1.99	1.71	1.59
Expected limit (pb)	3.23	2.77	2.45	2.37	2.17	1.98

#### D. Cross Section Limit

For inclusive double  $b$ -tagged dielectron and dimuon samples, Table IV shows the number of events of data, total expected background (from simulation and QCD), and expected signal in each invariant mass search window derived above. No significant excess is observed in any mass window, and hence the dijet invariant mass distributions are used to calculate the  $ZH$  cross section limits, with dielectron and dimuon samples combined. Limits are calculated using a modified frequentist approach [9]. The 95% C.L. limit is defined as the cross section at which the ratio of the confidence level for the sum of signal and background hypothesis,  $CL_{S+B}$ , to the confidence level for the background to represent the data,  $CL_B$ , reaches 0.05. The branching-ratio of  $Z \rightarrow e^+e^-$  or  $\mu^+\mu^-$  is 0.03366 [10]. Table V lists the expected and observed  $ZH$  cross section limits for each Higgs mass, derived from the combination of dielectron and dimuon samples. Figure 4 show the expected and observed  $ZH$  cross section limits, compared to the Standard Model expectation. The CDF limits showed in the same figure are the results by using Neural Net.

#### Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF

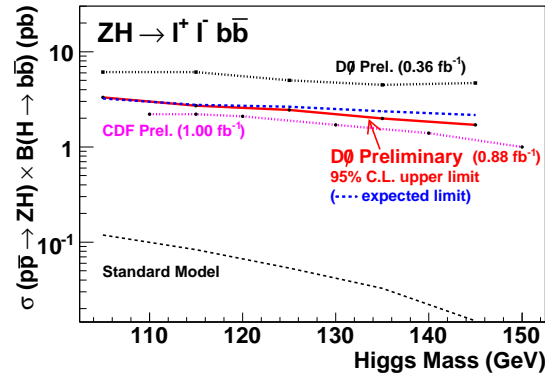


FIG. 4: The 95% confidence level upper limits on  $ZH$  cross sections. The limits are derived from the combination of  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$  samples. The CDF  $1 \text{ fb}^{-1}$  results used Neural Net.

(Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); Research Corporation, Alexander von Humboldt Foundation, and the Marie Curie Program.

- 
- [1] DØ Collaboration, V. Abazov *et al.*, “The Upgraded DØ Detector,” accepted to Nucl. Instrum. Methods Phys. Res. A.
  - [2] S. Abachi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994).
  - [3] G. C. Blazey *et al.*, in *Proceedings of the Workshop: “QCD and Weak Boson Physics in Run II,”* edited by U. Baur, R. K. Ellis, and D. Zeppenfeld, (Fermilab, Batavia, IL, 2000) p. 47; see Sec. 3.5 for details.
  - [4] Thomas Gadfort, Gavin Hesketh, Vincent Lesne, Mark Owen, Raimund Strohmer, Boris Tuchming, “Muon Identification Certification for p17 data”, DØNote 5157, Sept., 2006.
  - [5] Miruna Anastasoae, Stephen Robinson, and Tim Scanlon, “Performance of the NN *b*-tagging Tool on p17 Data”, DØnote 5213.
  - [6] J. Pumplin *et al.*, J. High Energy Phys. 07 (2002) 012.
  - [7] J. Campbell and K. Ellis, MCFM, *Monte Carlo for FeMtobarn processes*, <http://mcfm.fnal.gov/>.
  - [8] J. Heinmiller *et al.*, “A Search for *ZH* Production in *p $\bar{p}$*  Collisions at  $\sqrt{s} = 1.96$  TeV with the Full Pass2 Data Set”, DØnote 5185, June 6, 2006.
  - [9] T. Junk, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999); A. L. Read, Workshop on Confidence Limits, CERN-OPEN-2000-205 (2000).
  - [10] W.-M. Yao *et al.* (Particle Data Group), J. Phys. G **33**, 1 (2006).